In: A. Mullery, M. Besson, M. Campolargo, R. Gobbi and R. Reed, eds.: Intelligence in Services and Networks: Technology for Cooperative Competition. Berlin Heidelberg: Springer, 459-467.

# Personalized Hypermedia Information Provision through Adaptive and Adaptable System Features:

## User Modeling, Privacy and Security Issues

**Josef Fink, Alfred Kobsa, Jörg Schreck**
GMD – German National Research Center for Information Technology
Institute for Applied Information Technology (FIT)
Human-Computer Interaction Research Division (HCI)
D-53754 Sankt Augustin
+49 2241 14 {2729, 2315, 2859}
{josef.fink, alfred.kobsa, joerg.schreck}@gmd.de

## Abstract

Users of publicly accessible information systems are generally heterogeneous and have different needs. The aim of the AVANTI project is to cater to these individual needs by adapting the user interface and the content and presentation of WWW pages to each individual user. The special needs of elderly and handicapped users are also partly considered. A model of the characteristics of user groups and individual users, a model of the usage characteristics of the system, and a domain model are exploited in the adaptation process. This paper describes the detected differing needs of AVANTI users, the kind of adaptations that are currently implemented to cater to these needs, and the system architecture that enables AVANTI to generate user-adapted web pages from distributed multimedia databases. Special attention is given to privacy and security issues which are crucial when personal information about users is at stake.

## Keywords

Adaptive hypermedia, individualization, personalization, disabled and elderly users, adaptivity, adaptability, user modeling, user model server, privacy, security

## Introduction

The aim of AVANTI [1], a collaborative R&D project partially funded by the European Commission within the ACTS programme, is to develop and evaluate a distributed information system which provides hypermedia information about a metropolitan area (e.g., about public services, transportation, buildings) for a variety of users with different needs (e.g., tourists, citizens, travel agency clerks, elderly people, blind persons, wheelchair-bound people, and users with (slight) forms of dystrophy).

In order to develop an information service which is able to take the aims, interests, experiences, and abilities of its different users into account, AVANTI will take advantage of:

- methods and tools developed in the context of adaptive and adaptable systems during the last few years,
- standardized software components in the area of the World-Wide Web (WWW), and

- the widespread availability of computers interconnected in metropolitan-area networks.

The AVANTI system can be accessed from offices, public information booths, people's homes, and appropriate mobile computing devices (e.g. message pads and palmtops) throughout the world. Internal models of both user groups and individual users will help adapt the content and presentation to each user's individual needs.

## User Needs

Our investigations of the AVANTI user groups have shown that their needs are considerably heterogeneous. Moreover, individual differences in user needs have also been encountered. Some examples might illustrate this:

- For users who have never used the AVANTI system before, the topography of the hypermedia space should be kept simple (e.g., restricted to a sequence, grid, or tree [25]) in order to reduce the efforts necessary for building an appropriate mental model [7] [24]. Likewise, links to other hypermedia pages should be augmented by a label, or a short comment. Both adaptations can, however, be redundant (or even cumbersome) for citizens who use the information system of their home town frequently.
- For users interested in a specific subject, interesting details should be provided, e.g. an assessment of each painter in a web-based virtual museum. If the user lacks this specific interest, such detailed information should not be presented in order to reduce the efforts for building a mental model of the current hypermedia page [19] [24].
- For laypersons like tourists in a travel booking scenario, a technical term like 'check-out time' should be supplemented by an explanation. This is normally not necessary for domain experts like travel agents.
- For users with low-bandwidth network access (e.g., via a slow modem), information that requires high bandwidth (like videos and high-resolution pictures) should be replaced by less demanding but nevertheless appropriate equivalents.
- For blind users, the modality of the presented information must be changed in the case of tactile and/or audio output. Moreover, additional orientation and navigation aids (e.g., table of contents, indices) are helpful for this user group [14].
- For wheelchair-bound users, information concerning the accessibility of premises (e.g., the availability and the dimensions of ramps and elevators, the type and width of doors) is important and should therefore be provided.
- For users with (slight) forms of dystrophy, the man-machine interface (i.e., the interaction objects and associated manipulation techniques) should be adapted accordingly, i.e., should be made less sensitive to erratic hand movements.

When implementation issues are considered, it becomes obvious that all these needs can hardly be addressed within the scope of a single project. Consequently, we focused the further investigation on mainly mobility-related user requirements in the metropolitan areas of Siena (Italy), Rome (Italy), and Kuusamo (Finland) and consolidated the findings.

## Scope of Adaptivity and Adaptability

In order to cater to different user needs, information systems can be tailored manually by the user or automatically by the system. Systems that allow the user to change certain system parameters, and adapt their behavior accordingly, are called adaptable [20]. Systems that adapt to users automatically based on their assumptions about them are called adaptive.

Both features, adaptivity and adaptability, will be provided by the AVANTI system:

- *Adaptivity and adaptability within the user interface*
  We integrate and implement (special) I/O devices (e.g., macro mouse, Braille display, speech synthesizer), visual and non-visual interface objects, and associated interaction techniques [23].
- *Adaptivity and adaptability within hypermedia pages*
  We implement the adaptation of the information content, information modality, information prominence, orientation and navigation aids, search facilities, and links to other hypermedia pages [4].

Whereas the first group of adaptations aims at enabling and improving the overall access to the information system, the second group of adaptations aims at individualizing one specific hypermedia system.

## User and Usage Modeling

In order to provide user-oriented adaptivity, a so-called 'user model' has to be set up and maintained by the AVANTI system. A user model contains explicitly modeled assumptions which represent relevant characteristics of an individual user, like preferences and interests, domain knowledge, physical, sensorial, and cognitive abilities. Different methods for acquiring assumptions about the user have been discussed in the literature [6].

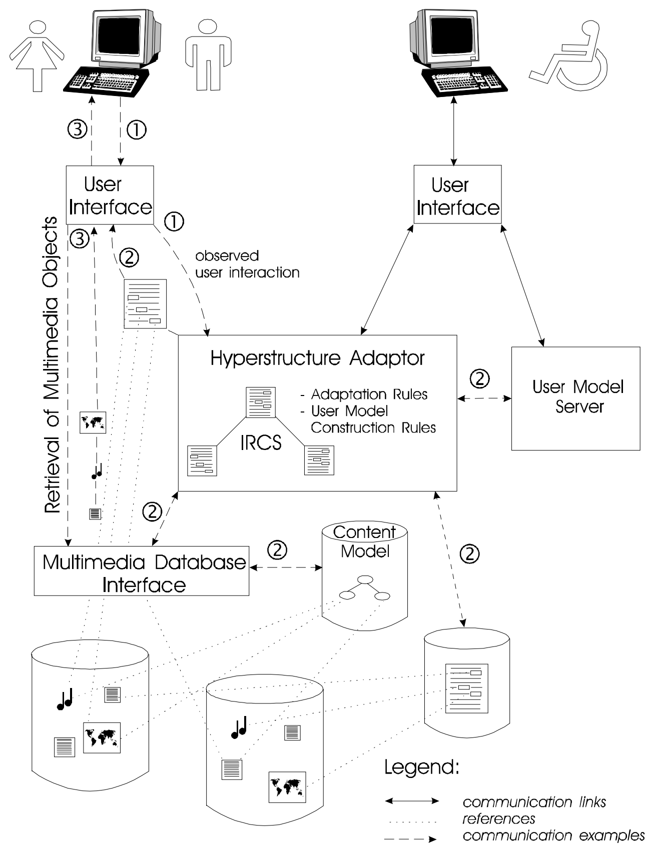
In AVANTI, assumptions will be acquired from the following sources of information:

- An initial interview provides the basis for primary assumptions about the user and is therefore a valuable source of information for initially assigning the user to certain user subgroups (see the 'stereotypes' below).
- Certain user actions can be exploited for the acquisition of primary assumptions. For instance, if the user requests an explanation for a technical term then he or she can be assumed not to be familiar with it [15].
- Based on primary assumptions about the user and additional information about the application domain, the system can draw inferences in order to acquire further assumptions about the user. For instance, if the user is interested in paintings, and being a tourist in Florence, has a special interest in the famous 'Galleria degli Uffizi' in Florence, we can predict the user's interest in Botticelli's 'The Birth of Venus'.
- So-called 'stereotypes' [22] contain assumptions about relevant characteristics of user subgroups (e.g., tourists, blind users). If certain preconditions are met, a stereotype can be activated for a specific user which means that the assumptions contained in the stereotype become assigned to the user.

In order to support technically motivated or usage-oriented adaptivity, a sub-component of the user model, the so-called ‘usage model’, contains relevant characteristics of the environment (e.g., terminal location, user interface characteristics) and the user’s interaction with the AVANTI system (e.g., history of visited pages, frequently requested pages, most likely future hypermedia page requests). Apart from information that is available *a priori*, such as about the environment of a specific terminal, most information in the usage model is elicited at run-time, either directly from hypermedia page requests via the HTTP [3] protocol or indirectly by employing statistical methods like regression analysis.

## System Architecture

The following figure shows the architecture of the AVANTI system:




In the following, we will focus on the functionality of, and the cooperation between, the main architectural components of the AVANTI system, namely the *User Interface* (UI), the *Hyperstructure Adaptor* (HSA), the *User Model Server* (UMS), and the *Multimedia Database Interface* (MDI) within the scenario of a request for a hypermedia page. The numbers refer to those in the figure:

① The user requests a hypermedia page. The UI forwards this request to the HSA.

② The HSA fetches the requested hypermedia page from secondary storage. The markup language used within this page is a subset of and an extension to HTML [28] named 'Information Resource Control Structure' (IRCS). Apart from static elements, an IRCS page may contain optional and alternative hypermedia objects, and also groups of hypermedia objects with an associated layout like a page header, toolbar, etc. An example for an optional element is supplementary information on wheelchair accessibility. Examples for alternative elements are technical vs. non-technical descriptions and an image of a painting vs. its textual description.
The processing of these optional and alternative elements is controlled by *Adaptation Rules*, which can take information from other system components into account, namely assumptions about user characteristics (e.g., knowledge, interests, preferences) from the UMS, and content-related information about multimedia objects from the *Content Model* (CM) via the MDI. Information about the current user's session (e.g., previously requested IRCS pages, previously provided input) is available as well. A second group of rules that may be contained in this IRCS page are *User Model Construction Rules*. They control the formation of so-called primary assumptions about the user (i.e., assumptions which are directly derived from the user's interaction with the hypermedia page). Primary assumptions are directly reported from the HSA to the UMS.
The HSA interprets the requested hypermedia page and the Adaptation Rules, generates an adapted page (which is compliant to standard HTML) and hands it over to the UI for presentation.

③ The UI interprets the hypermedia page, retrieves multimedia objects from the AVANTI databases transparently via the MDI, and finally presents the requested hypermedia page to the user. [1]

The communication between all active components is carried out via the HTTP protocol. On top of it, a restricted and slightly enhanced version of KQML (Knowledge Query and Manipulation Language [8]) for user modeling purposes is used for communication with the UMS [17].

The main advantages of this architecture include the following:

- Already existing software in the area of the WWW (e.g., communication libraries, browsers, servers, proxies, web development environments, and database gateways) can partially or fully be used for the development of AVANTI components. This allows the developers to focus on adaptivity and adaptability, and on the evaluation of these concepts in several field tests.
- Most WWW browsers available today can access the AVANTI system and take advantage of the customization features that are based on user-oriented adaptivity and adaptability.

---

[1] As pointed out before, the UI is able to perform additional adaptations (e.g., use alternative I/O devices, visual and non-visual interface objects and associated interaction techniques) which are not further discussed here.

- All active components within AVANTI can be fully distributed according to organizational and technical requirements. This is achieved by employing an HTTP-based name service for resolving symbolic references at run-time [13].
- Certain content adaptations may be dynamically delegated from the HSA to the UI, if the necessary environment for the execution of (mobile) Java code [5] is present there. Delegated adaptations relieve the server-based HSA and allow for a more scalable architecture, avoiding the inherent limitations of a purely server-based approach.

The HSA and the UMS are central constituents of the AVANTI architecture. Their development does not have to be started from scratch since already available software can be employed as a basis, including 'WebObjects' [18] for the HSA and 'BGP-MS' (Belief, Goal and Plan Maintenance System [16]) for the UMS.

## Security and Privacy Issues

The distributed architecture of AVANTI implies that its constituents communicate via network connections. The distribution and the fact that the system is shared between the user and the information provider poses challenges regarding the security and privacy of the users being modeled. In the following, we briefly discuss technical means for ensuring secure and private communication between the constituents of the AVANTI architecture. Moreover, we will outline various user modeling policies.

Encryption techniques provide the basis for secure information flow. *Link encryption* based on the hardware of network connections (i.e., on layer 1 or 2 of the ISO/OSI reference model [12]) provides a simple and transparent means for keeping transport data private (see e.g. the proposed Internet Protocol standard 'IPv6' [11]). In order to take advantage of link encryption as an end-to-end service, all physical nodes within a communication channel have to support this kind of service. At the moment, this requirement is not generally met.

Software solutions offer various opportunities to shield data transparently on the transport layer (i.e., on layer 3 and 4 of the ISO/OSI reference model). For example, particular implementations of TCP/IP establish protected communication channels. In order to take advantage of these encryption services on the transport level, all (potentially heterogeneous) operating systems that are hosting AVANTI components must employ compatible implementations of secure TCP/IP. Up to now, this precondition is normally not met.

Recent developments like 'Secure Socket Layer' (SSL, [10]) and 'Personal Communication Technologies' (PCT, [2]) reside above the transport layer and allow for safe communication between applications. The only requirement that has to be met is that the communicating parties actually use SSL (or PCT). This precondition is normally met since these implementations are available on many platforms and are interoperable. On the presentation layer (i.e., layer 6 of the ISO/OSI reference model), high-level protocols can be employed in order to tailor the security mechanisms to the respective needs and to realize *end-to-end encryption*. Secure HTTP (S-HTTP, [21]) allows for different modes of protection depending on the kind of transported data. The 'Protocol Extension Protocol' (PEP, [27]) and the 'Security Extension Architecture' (SEA, [26]) of the World-Wide Web Consortium offer mechanisms to communicate that transportation security and authentication is required. A certification authority has to be added to the overall infrastructure in order to authenticate the communicating parties.

For the AVANTI system, a dual approach is appropriate. S-HTTP is recommended for the safe exchange of hypermedia pages between the HSA and the UI. This would enable end-to-end encryption and authentication between these components using already established WWW standards. The KQML-based communication between the UMS, the HSA, and the UI should also meet these security requirements. An extension to KQML like the one proposed in [9] would allow the security aspects (i.e., encryption techniques) to be negotiated within a communication that can take the sensitivity of the transported data into account.

Confidentiality must not only be guaranteed for data exchange but also for data storage since personal information about users resides in the UMS. These data include usage records with time stamps, data that the user supplied, and assumptions that were inferred from the user's data and usage behavior. Privacy issues arise if a user accesses the system by revealing his or her identity rather than remaining anonymous. If the user provides information on disabilities and interests, this data is not only person-related but possibly even sensitive. Several options should be offered by the system in order to accommodate user's privacy expectations:

- If possible, users should be given the option of accessing the system anonymously (e.g., with a pseudonym) if they do not want to reveal their identities (such an option is even an obligatory provision in the current draft of the forthcoming German multimedia law.).
- In an (optional) initial dialog the user should be able to choose between
    - no user modeling,
    - short-term user modeling, e.g. for the current session only,
    - long-term modeling using persistent user models that are augmented with information from the current session.
- At the end of each session the user should be asked if his or her model is to be deleted or stored for subsequent sessions.

These measures are taken to meet legal regulations regarding systems that process personal information and to increase user acceptance by making the system transparent. The fact that data about the user are gathered and processed should be pointed out to the user at the beginning of each session.

## Related Projects

The main motivation for developing adaptive hypertext and hypermedia systems is the overwhelming growth of many hypermedia spaces (e.g., the WWW) in terms of size, complexity, and heterogeneity. Likewise, the user population which is confronted with these hypermedia spaces is growing, also both in size and heterogeneity. In order to keep pace with these trends, at least twenty adaptive hypertext and hypermedia systems have been developed in the last few years in order to provide more sophisticated tools for orientation, navigation, and search (for an overview and a brief description of most of these systems we refer to Brusilovsky [4]). While AVANTI shares characteristics with some of them, there are also several important distinguishing features including the following:

- In AVANTI, user-oriented *adaptations* take place *within hypermedia pages* as well as *within the user interface*. This seems to be especially beneficial for disabled users (e.g., blind users), since they can often take advantage of adaptations on both levels. The extension of hypermedia adaptation techniques for catering also to perceptual and motoric disabilities is a unique characteristic of AVANTI.
- AVANTI puts an *extendible set of adaptation techniques*, namely adaptive presentation of multimedia elements, direct guidance, adaptive sorting, hiding, and annotation of links at the disposal of the hypermedia author.
- The *adaptations* offered in AVANTI *address heterogeneous user needs*, *usage patterns*, and *environmental conditions* (e.g., technical capabilities of the user interface, available network bandwidth), and provide therefore a more holistic approach to the challenges mentioned at the beginning of this chapter.
- In AVANTI, user models are entirely hosted by a *central user model server* which offers the user and other AVANTI components location-independent access to the most recent user-related information. Moreover, synergetic effects with respect to user-related information can be expected (e.g., the HSA takes advantage of assumptions acquired by the UI and vice versa). Another advantage of this centralized user modeling approach is that the other components of the AVANTI system become totally relieved of user modeling tasks and can take advantage of sophisticated run-time services of the UMS.

## References


1. AVANTI Home Page. Available at http:// www.gmd.de/ fit/ projects/ avanti.html
2. Benaloh, J.; Lampson, B.; Simon, D.; Spies, T.; Yee, B.: Personal Communication Technologies, Microsoft Corporation, 1995. Available at http:// www.lne.com/ ericm/ pct.html
3. Berners-Lee, T.; Fielding, R.; Frystyk, H.: Request for Comments 1945, Hypertext Transfer Protocol - HTTP/1.0, Category: Informational, 1996. Available at http:// www.csl.sony.co.jp/ cgi-bin/ hyperrfc?1945
4. Brusilovsky, P.: Methods and Techniques of Adaptive Hypermedia, User Modeling and User-Adapted Interaction 6(2-3), 1996, pp. 87-129.
5. Campione, M.; Walrath, K.: The Java Tutorial, 1996. Available at http:// www.javasoft.com/ tutorial/ index.html
6. Chin, D. N.: Acquiring User Models, Artificial Intelligence Review 7, pp. 185-197, 1993.
7. Conklin, J.: Hypertext: An Introduction and Survey, IEEE Computer, September 1987, pp. 17-41.
8. Finin, T. W.; Weber, J.; Widerhold, G.; Genesereth, M.; Fritzson, R.; McKay, D.; McGuire, J.; Pelavin, R.; Shapiro, S.; Beck, C.: Specification of the KQML Agent-

Communication Language, 1993. Available at http:// www.cs.umbc.edu/ kqml/ papers/ kqmlspec.ps

9. Finin, T. W.; Mayfield, J.; Thirunavukkarasu, C.: Secret Agents - A Security Architecture for the KQML Agent Communication Language, Intelligent Information Agents Workshop (CIKM'95), Baltimore, 1995.

10. Freier, A. O.; Karlton, P.; Kocher, P. C.: The SSL Protocol, Version 3.0, 1996. Available at ftp:// ietf.org/ internet-drafts/ draft-ietf-tls-ssl-version3-00.txt

11. Huitema, C.: IPv6 - The New Internet Protocol, Prentice-Hall, Englewood Cliffs, New Jersey, 1996.

12. Hunt, C.: TCP/IP Network Administration, O'Reilly & Associates, Sebastopol, California, 1992.

13. KAPI, 1996. Available at http:// hitchhiker.space.lockheed.com/ pub/ aic/ shade/ software/ KAPI/

14. Kennel, A.; Perrochon, L.; Darvishi, A.: WAB: World-Wide Web Access for Blind And Visually Impaired Computer Users. New Technologies in the Education of the Visually Handicapped, Paris, June 1996 and ACM SIGCAPH Bulletin, June 1996. Available at http:// www.inf.ethz.ch/ department/ IS/ ea/ blinds/

15. Kobsa, A.; Müller, D.; Nill, A.: KN-AHS: An Adaptive Hypertext Client of the User Modeling System BGP-MS, Proceedings of the Fourth International Conference on User Modeling, Hyannis, MA, pp. 99-105, 1994.

16. Kobsa, A.; Pohl, W.: The User Modeling Shell System BGP-MS, User Modeling and User-Adapted Interaction 4(2), pp. 59-106, 1995.

17. Kobsa, A.; Fink, J.; Pohl, W.: A Standard for the Performatives in the Communication between Applications and User Modeling Systems (draft), 1996. Available at ftp:// ftp.informatik.uni-essen.de/ pub/ UMUAI/ others/ rfc.ps

18. NeXT Corporation: WebObjects, 1996. Available at http:// www.next.com/ WebObjects/ Products.html

19. Nielsen, J.: The Art of Navigating through Hypertext, Communications of the ACM, vol. 33, no. 3 (March 1990), pp. 296-310.

20. Oppermann, R. (Ed.): Adaptive User Support - Ergonomic Design of Manually and Automatically Adaptable Software, Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1994.

21. Rescorla, E.; Schiffman, A.: The Secure HyperText Transfer Protocol, Enterprise Integration Technologies, 1995. Available at http:// www.eit.com/ creations/ s-http/ draft-ietf-wts-shttp-00.txt

22. Rich, E.: User Modeling via Stereotypes, Cognitive Science, 3, pp. 329-354, 1979.

23. Savidis, A.; Stephanidis, C.: Developing Dual User Interfaces for Integrating Blind and Sighted Users: the HOMER UIMS, Proceedings of the CHI'95 Conference on Human Factors in Computing Systems, Denver, Colorado, May 7-11, 1995.

24. Schaumburg, H.; Issing, L. J.: Lernen mit Hypermedia: Verloren im Hyperraum?, HMD - Theorie und Praxis der Wirtschaftsinformatik, No. 190, pp. 108-121, 1996.
25. White, B.: Web Document Engineering, Tutorial Notes, Fifth International World-Wide Web Conference, Paris, May 1996, O'Reilly & Associates, Sebastopol, California, 1996.
26. World Wide Web Consortium: SEA: A Security Extension Architecture for HTTP/1.x, W3C Working Draft, 1996. Available at http:// www.w3.org/ pub/ WWW/ TR/ WD-http-sea.html
27. World Wide Web Consortium: HTTP/1.2 Extension Protocol (PEP), W3C Working Draft, 1996. Available at http:// www.w3.org/ pub/ WWW/ TR/ WD-http-pep.html
28. World Wide Web Consortium: HyperText Markup Language (HTML), 1996. Available at http:// www.w3.org/ pub/ WWW/ MarkUp/